RNA (1999), *5*:195–205. Cambridge University Press. Printed in the USA.

RNA binding by the novel helical domain of the influenza virus NS1 protein requires its dimer structure and a small number of specific basic amino acids

WEIRONG WANG,[1] KELLY RIEDEL,[1] PATRICIA LYNCH,[1] CHEN-YA CHIEN,[1,2] GAETANO T. MONTELIONE,[1,2] and ROBERT M. KRUG[1]

[1]Department of Molecular Biology and Biochemistry, Rutgers University, Piscataway, New Jersey 08855-1179, USA
[2]Center for Advanced Biotechnology and Medicine, Piscataway, New Jersey 08855-1179, USA

ABSTRACT

The RNA-binding/dimerization domain of the NS1 protein of influenza A virus (73 amino acids in length) exhibits a novel dimeric six-helical fold. It is not known how this domain binds to its specific RNA targets, one of which is double-stranded RNA. To elucidate the mode of RNA binding, we introduced single alanine replacements into the NS1 RNA-binding domain at specific positions in the three-dimensional structure. Our results indicate that the dimer structure is essential for RNA binding, because any alanine replacement that causes disruption of the dimer also leads to the loss of RNA-binding activity. Surprisingly, the arginine side chain at position 38, which is in the second helix of each monomer, is the only amino-acid side chain that is absolutely required only for RNA binding and not for dimerization, indicating that this side chain probably interacts directly with the RNA target. This interaction is primarily electrostatic, because replacement of this arginine with lysine had no effect on RNA binding. A second basic amino acid, the lysine at position 41, which is also in helix 2, makes a strong contribution to the affinity of binding. We conclude that helix 2 and helix 2′, which are antiparallel and next to each other in the dimer conformation, constitute the interaction face between the NS1 RNA-binding domain and its RNA targets, and that the arginine side chain at position 38 and possibly the lysine side chain at position 41 in each of these antiparallel helices contact the phosphate backbone of the RNA target.

Keywords: Influenza virus NS1 protein; protein dimerization; RNA-binding domain; RNA-binding epitope

INTRODUCTION

The NS1 protein of influenza A virus (NS1A protein) is an RNA-binding protein with multiple functions. Specifically, it inhibits the nuclear export of mRNAs that contain 3′ poly(A) ends (Alonso-Caplen et al., 1992; Fortes et al., 1994; Qian et al., 1994; Qiu & Krug, 1994), inhibits pre-mRNA splicing (Fortes et al., 1994; Lu et al., 1994; Qiu et al., 1995; Wang & Krug, 1998), and blocks the activation of the dsRNA-activated protein kinase that would result in the inhibition of translation (Hatada & Fukuda, 1992; Lu et al., 1995). The functional domains of the NS1A protein were identified by mutagenesis experiments in which groups of two or three amino acids at various positions along the protein chain were replaced with alanines (Qian et al., 1994). The RNA-binding domain, which is at the amino-terminus (Qian et al., 1994), has four known targets: double-stranded RNA (dsRNA), U6 snRNA, U6atac snRNA, and poly(A) (Hatada & Futada, 1992; Qiu & Krug, 1994; Lu et al., 1994, 1995; Qiu et al., 1995; Wang & Krug, 1998). An amino-terminal fragment, which is comprised of the first 73 amino-terminal amino acids (NS1A(1–73)), possesses all the RNA-binding properties of the full-length protein (Qian et al., 1995). Both the NS1A protein and the NS1A(1–73) fragment have been shown to exist as dimers in the absence of their RNA targets and when they are bound to U6 snRNA (Nemeroff et al., 1995; Qian et al., 1995). Mutagenesis experiments suggested that the RNA-binding and dimerization domains are coincident (Nemeroff et al., 1995). The other functional domain, the effector domain, which is located in the carboxy-terminal half of the molecule, was proposed to interact with host nuclear proteins to carry out several

Reprint requests to: Robert M. Krug, Department of Molecular Biology, Rutgers University, CABM-305, 679 Hoes Lane, Piscataway, New Jersey 08855-1179, USA; e-mail: krug@waksman.rutgers.edu.

functions of the NS1A protein in vivo (Qian et al., 1994). This proposal has been verified by the demonstration that the NS1A effector domain functionally interacts with the cellular 30 kDa subunit of the cleavage and polyadenylation specificity factor, an essential component of the 3′ end-processing machinery of cellular pre-mRNAs (Nemeroff et al., 1998).

The three-dimensional structure of the RNA-binding/dimerization domain of the NS1 protein of influenza A/Udorn/72 virus has been determined by both nuclear magnetic resonance (NMR) and X-ray crystallography (Chien et al., 1997; Liu et al., 1997). This domain is almost totally α-helical and exhibits a novel dimeric six-helical chain fold (Chien et al., 1997; Liu et al., 1997). Although several α-helical RNA-binding domains have been described (Banner et al., 1987; Stebbins et al., 1995; Battiste et al., 1996; Berglund et al., 1997; Hinck et al., 1997; Xing et al., 1997; Davies et al., 1998; Markus et al., 1998), β-sheets are the predominant secondary structures in most RNA-binding proteins whose structures have been determined to date (reviewed in Nagai, 1996). One α-helical RNA-binding domain that has been characterized is that of the Rev protein of human immunodeficiency virus-1 (HIV-1) (Battiste et al., 1996). A short α-helical peptide corresponding to the sequence of the Rev RNA-binding domain was shown to bind in, and distort, the major groove of its RNA target. In addition, RNA binding by the L11 ribosomal protein has been shown to be mediated largely by its α-helical surface (Hinck et al., 1997; Xing et al., 1997).

It is not known how the six-helical RNA-binding domain of the NS1A protein binds to its RNA targets. As an initial approach to elucidating the mode of this RNA binding, we have introduced single alanine replacements into the NS1A RNA-binding domain at specific positions in the three-dimensional structure. This approach enabled us to identify the specific basic amino acids that participate only in RNA binding and not in dimerization. Consequently, we conclude that these amino acids interact with the RNA per se, thereby defining the RNA-binding epitope. In addition, we show that the dimer structure is essential for RNA binding, because any alanine replacement that causes disruption of the dimer results in the loss of RNA-binding activity.

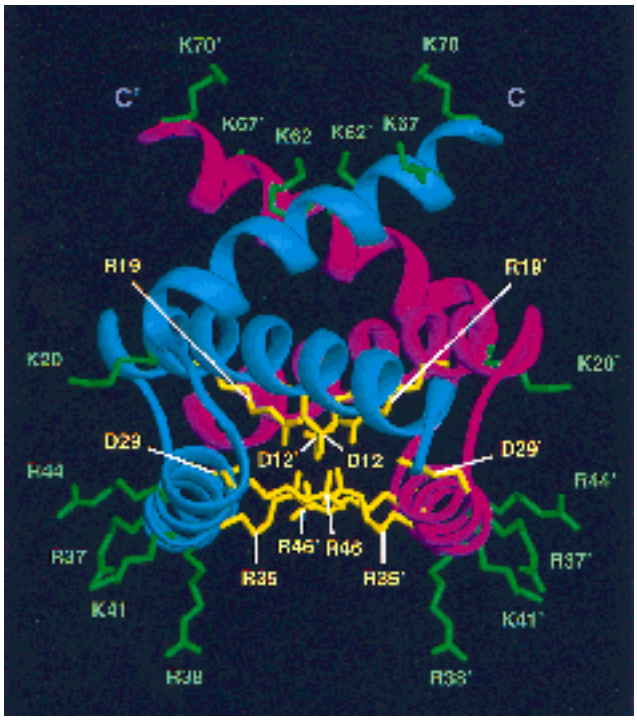


FIGURE 1. The three-dimensional structure of the NS1A(1–73) polypeptide (pdb entry 1ns1). The solvent-exposed basic amino acids are shown in green, and the amino acids that are expected to stabilize the dimer structure are shown in yellow.

RESULTS

Dimerization is required for RNA-binding activity

The NS1A(1–73) RNA-binding domain exists as a dimer in solution (Qian et al., 1995; Chien et al., 1997; Liu et al., 1997; Fig. 1). Each monomer is comprised of three α-helices: helix 1, residues 3–25; helix 2, residues 30–50; helix 3, residues 54–69. The three helices are connected by two short turns. In the dimer, the monomers are tightly packed together in a head-to-tail topology. To determine whether the dimer structure per se is required for RNA binding, five amino acids that would be predicted to be required for the maintenance of the dimer structure were individually replaced with an alanine residue (Fig. 1, these amino acids are shown in yellow): D12, R19, D29, R35, and R46, which form intermolecular hydrogen bonds and electrostatic interactions at the dimer interface (Chien et al., 1997; Liu et al., 1997) and contribute to a structured water network (Liu et al., 1997).

These mutant NS1A(1–73) proteins were first tested for their ability to bind to a short (55-bp) dsRNA using a gel shift assay (Fig. 2A). All of these mutant proteins have lost their ability to bind this dsRNA, as well as longer dsRNAs (data not shown). In addition, these mutant NS1A(1–73) proteins were tested for their ability to form dimers using the glutathione Sepharose selection assay described in Materials and Methods. In vitro-labeled wild-type, full-length NS1A protein was combined with each of the purified GST-tagged mutant NS1A(1–73) polypeptides, and the mixtures were selected on glutathione Sepharose beads. All of the five NS1A(1–73) mutant polypeptides listed above fail to form heterodimers with the wild-type NS1A protein (Fig. 2B), indicating that these amino-acid side chains are indeed required for dimer formation. The same results were obtained using a glutaraldehyde crosslink-

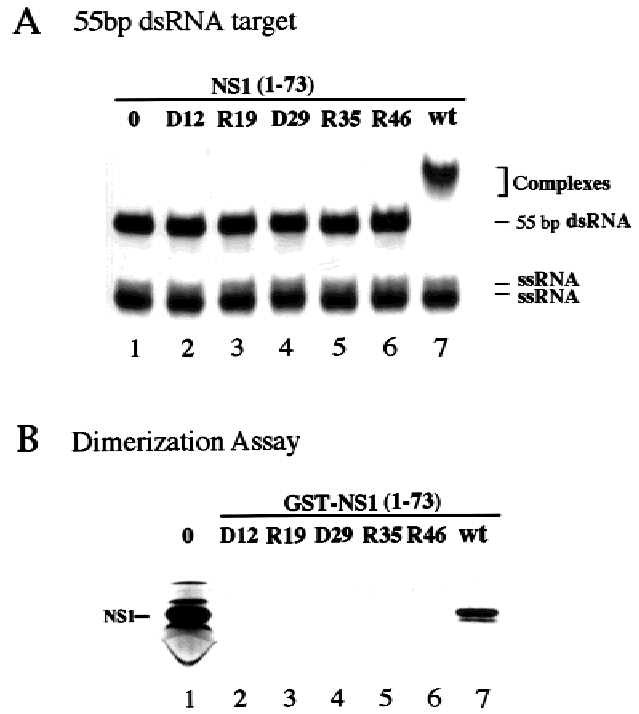


FIGURE 2. The NS1A(1–73) polypeptide can only bind to RNA as a dimer. The designation on top of each lane indicates the amino acid that has been replaced by alanine. **A**: Gel shift assay. The indicated mutant or wild-type polypeptide (400 nM) was incubated with dsRNA (10,000 cpm, 1 nM). The polypeptide–RNA complexes were separated from free RNA by nondenaturing gel electrophoresis. Lane 1: dsRNA alone; lane 7: wild-type NS1A(1–73). **B**: Glutathione Sepharose selection assays. Ten microliters of the in vitro-translated, ^{35}S-labeled, full-length NS1A protein was incubated with 20 μL glutathione Sepharose 4B beads and 5 μg of purified GST-NS1A(1–73) containing the indicated alanine replacement. After extensive washing, the ^{35}S-labeled, full-length NS1A protein that remained bound to the beads was eluted and analyzed by electrophoresis on a 14% SDS-PAGE gel. Lane 1: in vitro translated NS1A protein before selection.

ing assay (Nemeroff et al., 1995 and data not shown). Thus, the loss of RNA-binding activity for the D12A, R19A, D29A, R35A, and R46A mutant polypeptides can be attributed to the loss of intermolecular interactions required for dimerization. Accordingly, we conclude that only the dimeric form of NS1A(1–73) can bind to RNA.

Identification of the amino acids of the NS1A protein that are required for RNA binding, but not for protein dimerization

Basic amino acids on the surface of RNA-binding proteins are candidates for the residues that interact directly with the negatively charged phosphate backbone of RNA. Accordingly, we determined whether replacing specific basic amino acids of the NS1A(1–73) polypeptide with alanine caused the loss of dsRNA binding but not dimerization. The candidate basic amino acids were deduced from the structure of the dimeric NS1A(1–73) RNA-binding domain. Of the eleven basic amino acids in the NS1A(1–73) RNA-binding fragment, eight (K20, R37, R38, K41, R44, K62, K67, and K70; shown in green in Fig. 1) are completely, or partially, exposed to the solvent in the dimeric conformation and were considered to be possible candidates to interact with RNA per se.

Each of these basic amino acids was individually replaced with an alanine residue, and the resulting NS1A(1–73) molecules were tested for their ability to bind to a short (55-bp) dsRNA using a gel shift assay (Fig. 3A). Alanine replacements of basic amino acids in helix 3 (K62, K67, and K70) do not cause a loss in dsRNA binding (Fig. 3A, lanes 8–10). However, the RNA complex formed with the NS1A(1–73) molecule containing an alanine replacement at amino acid 62 has an increased mobility compared to the RNA complex formed with wild-type NS1A(1–73) (Fig. 3A, compare lanes 2 and 8). The altered mobility of the complex containing the mutant protein may indicate that it has a structure that differs from that of the complex containing the wild-type protein. The alanine replacement of amino acid K20 in helix 1 also does not cause a loss in dsRNA binding, although this protein–RNA complex also migrates faster than the complex containing wild-type NS1A(1–73) (Fig. 3A, lane 3). In contrast, two of the alanine replacements in helix 2 (R38 and K41) result in NS1A(1–73) molecules that have lost all detectable ability to shift the short dsRNA (Fig. 3A, lanes 5 and 6). The other two alanine replacements in helix 2 (R37 and R44) yield NS1A(1–73) molecules that exhibit only slight gel shifts of the dsRNA target (Fig. 3A, lanes 4 and 7).

The amount of either of these mutant polypeptides that is required to form these complexes is similar to the amount of wild-type needed to cause a larger shift of this dsRNA target (data not shown), suggesting that this slight shift is the result of protein–RNA interaction (see Fig. 5A and Discussion). As will be verified below, the small gel shifts observed for these complexes indicate that these NS1A(1–73) molecules do bind dsRNA, although these complexes have other changes in structure and/or electrostatics that result in a mobility that differs only slightly from free dsRNA. Similar results were obtained using U6 snRNA as the target (Fig. 3B); only the R38A and K41A mutant polypeptides exhibit a complete loss of RNA-binding activity. These mutant NS1A(1–73) molecules were also tested for their ability to form dimers (Fig. 3C). All of these eight mutant NS1A(1–73) polypeptides form heterodimers with the wild-type NS1A protein, indicating that these eight basic amino-acid side chains are not required for dimer formation.

To determine whether the same results are observed with a higher affinity RNA target, we used a longer dsRNA molecule (140 bp), which has about a tenfold higher affinity than the shorter (55-bp) dsRNA for the wild-type NS1A(1–73) RNA-binding domain (Fig. 4).

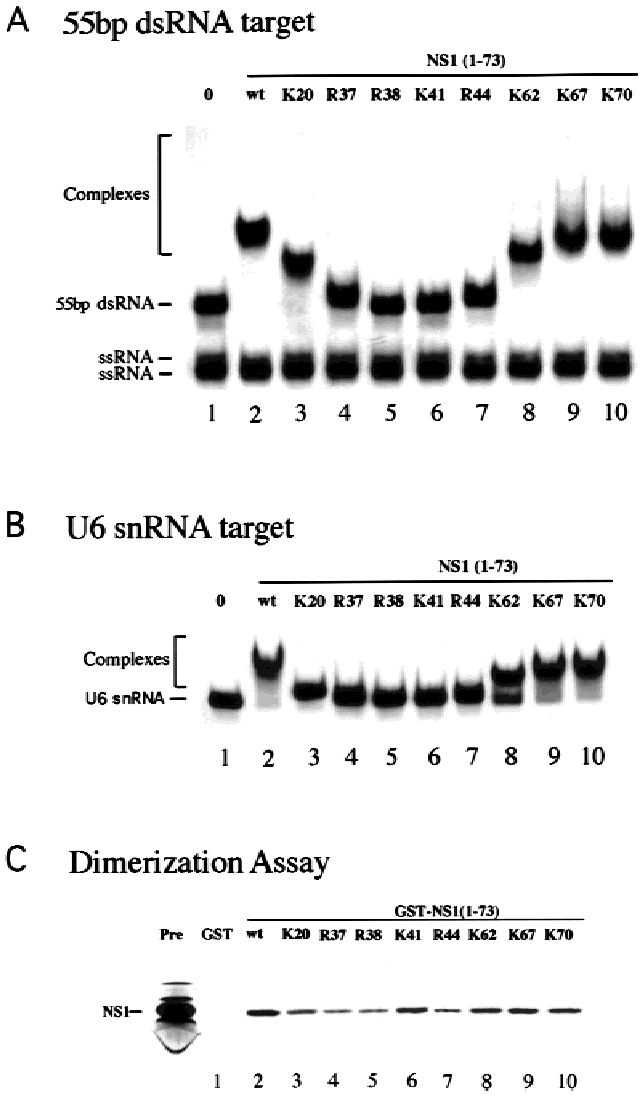


FIGURE 3. Identification of the basic amino acids of NS1A(1–73) that are required for RNA binding but not protein dimerization. The designation on top of each lane indicates the amino acid that has been replaced by alanine. **A**: Gel shift assay of the NS1A(1–73) mutant polypeptides using a 55-bp dsRNA as target. The indicated mutant or wild-type polypeptide (400 nM) was incubated with the 55-bp dsRNA (10,000 cpm, 1 nM), and the polypeptide–RNA complexes were separated from free RNA by nondenaturing gel electrophoresis. Lane 1: dsRNA alone; lane 2: wild-type NS1A(1–73) protein. **B**: Gel shift assays using U6 snRNA as the RNA target. **C**: Glutathione Sepharose selection assay for dimerization carried out as described in the legend to Figure 2. Lane 1: the in vitro-translated NS1A protein before selection; lane 2: nonfusion GST protein.

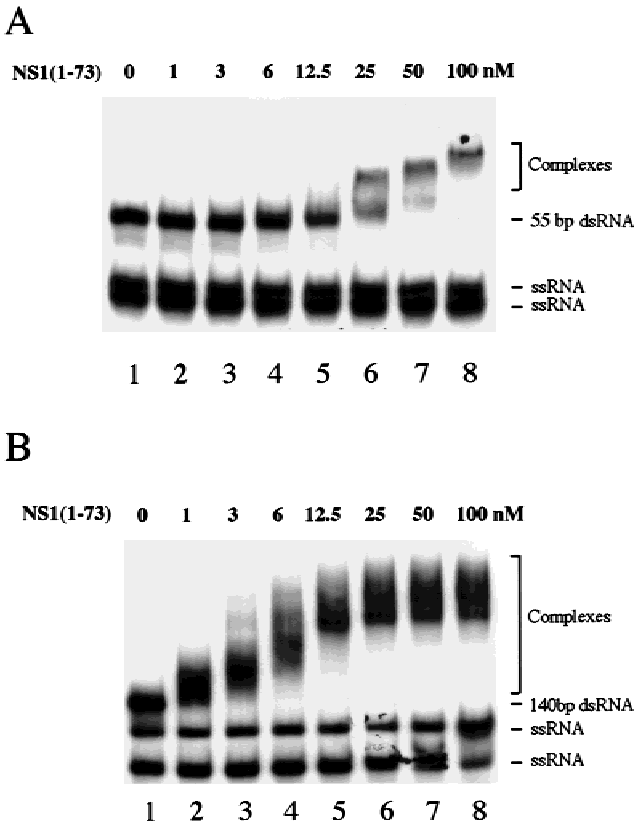


FIGURE 4. The affinity of the NS1A(1–73) polypeptide is higher for a longer dsRNA than for a shorter dsRNA. Increasing concentrations of wild-type NS1A(1–73) polypeptide were incubated with a 55-bp dsRNA (**A**) or 140-bp dsRNA (**B**). The polypeptide–RNA complexes were separated from free RNA by nondenaturing gel electrophoresis.

The binding of NS1A(1–73) to dsRNA is not sequence specific (Lu et al., 1995), and probably occurs at multiple sites along the dsRNA chain. Consequently, the higher affinity observed for longer-length dsRNA target molecules can be attributed at least in part to the increased configurational entropy resulting from the larger number of binding sites in the longer dsRNA target. The NS1A(1–73) molecule containing an alanine replacement at R38 does not bind the longer 140-bp dsRNA (Fig. 5A, lane 3), as was the case for the shorter dsRNA. In contrast, NS1A(1–73) molecules containing an alanine replacement at either R37 or R44 bind the longer dsRNA and form a complex with a mobility similar to that of the complex containing a wild-type NS1A(1–73) molecule (Fig. 5A, lanes 2 and 5). Thus, the basic side chains of R37 and R44 are indeed not required for RNA binding.

In addition, the K41A mutant polypeptide also forms a complex with the 140-bp dsRNA that has mobility similar to that formed with the wild-type polypeptide (Fig. 5A, lane 4), whereas the same concentration of this mutant protein does not detectably gel shift either the shorter (55-bp) dsRNA or U6 snRNA (compare lane 6 of Fig. 5A to Figs. 3A and 3B). However, as shown by the concentration dependence of RNA binding, the K41A mutant polypeptide has about a tenfold lower affinity than the wild-type molecule for the 140-bp dsRNA (Fig. 5B, compare lanes 7–11 with lanes 1–6). Based on these results, we conclude that the K41 side chain is not absolutely required for RNA binding, but makes a strong contribution to the affinity of binding, which is most evident when the RNA target has a lower affinity for the NS1A(1–73) molecule. In contrast to the K41A mutant, the R37A mutant polypeptide has essen-

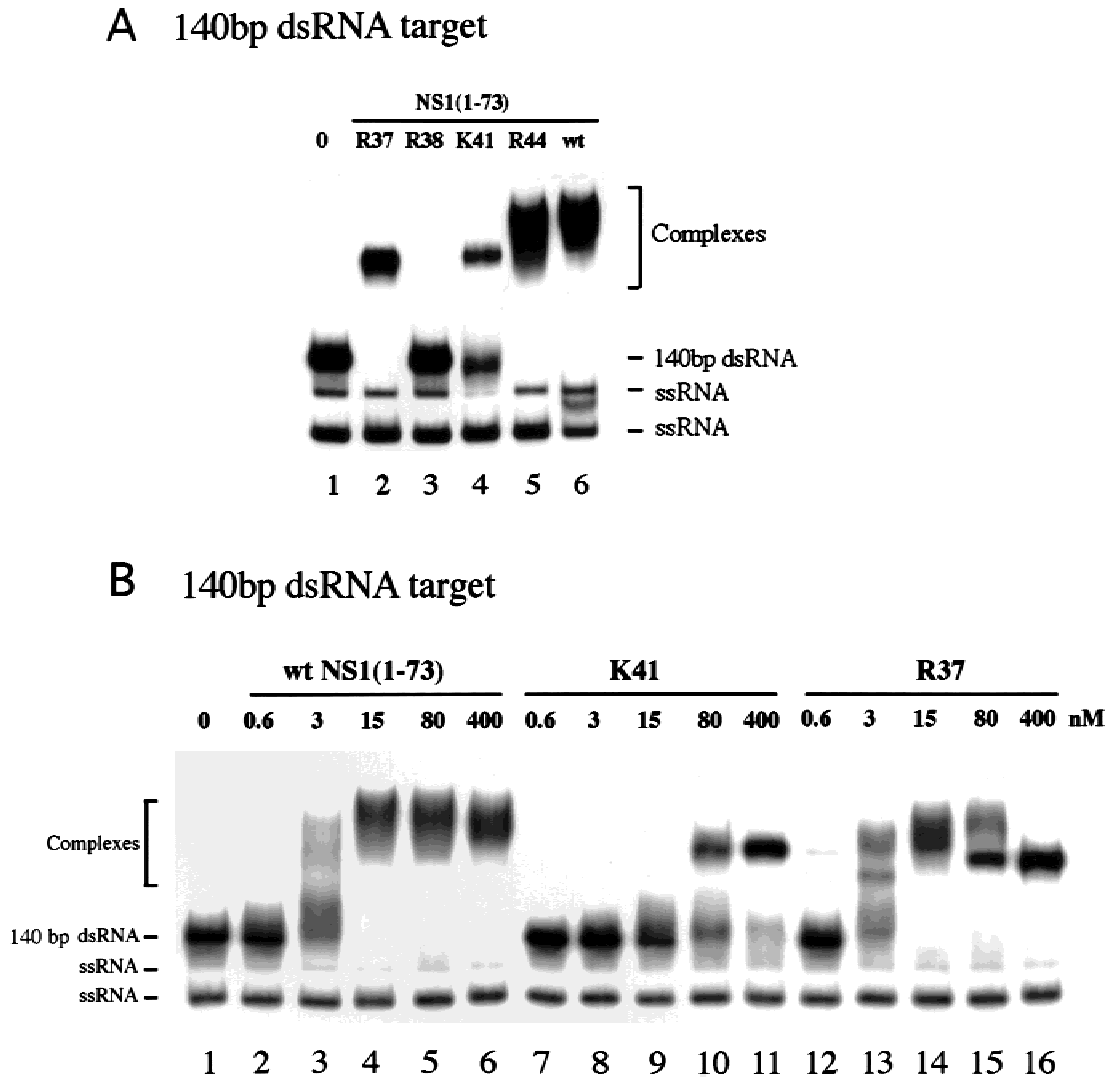


FIGURE 5. Identification of the basic amino acids of NS1A(1–73) that are required for binding a longer (140-bp) dsRNA. **A**: The indicated mutant or wild-type NS1A(1–73) polypeptide (400 nM) was incubated with the 140-bp dsRNA (10,000 cpm, 1 nM), and the polypeptide–RNA complexes were separated from free RNA by nondenaturing gel electrophoresis. Lane 1: dsRNA alone; lane 6: wild type NS1A(1–73) polypeptide. **B**: The NS1A(1–73) molecule containing an alanine replacement at K41, but not at R37, binds the 140-bp dsRNA with a lower affinity than wild-type NS1A(1–73). Increasing amounts of the wild-type (lanes 2–6), the K41 mutant (lanes 7–11), or the R37 mutant (lane 12–16) polypeptide were incubated with the 140-bp dsRNA target, and RNA binding was measured by gel shift assays. Lane 1: dsRNA alone.

tially the same affinity as the wild-type molecule for the 140-bp dsRNA, as shown at lower concentrations of these two polypeptides (Fig. 5B, compare lanes 12–14 with lanes 2–4). However, at higher concentrations of the R37A polypeptide, the complexes it forms with dsRNA have decreased mobilities compared to those formed at lower concentrations of this polypeptide (Fig. 5B, compare lanes 15 and 16 with lane 14), indicating that these complexes change in structure and/or electrostatics as the concentration of the R37A polypeptide is increased (see Discussion).

Consequently, our results with both the longer and shorter dsRNA molecules establish that the R38 side chain is the only one that is absolutely required for RNA binding, indicating that it probably interacts directly with the RNA target. To determine whether this interaction is based on electrostatic interactions or on specific hydrogen-bonding interactions, the arginine at position 38 was replaced with a lysine residue. Increasing amounts of R38K and wild-type NS1A(1–73) polypeptides were incubated with a 55-bp dsRNA target, and binding of dsRNA was measured by gel shift assay (Fig. 6). The R38K polypeptide binds to the RNA target with the same affinity as the wild-type protein, indicating that the interaction between RNA and the basic amino acid at position 38 of the NS1 RNA-binding domain is based on electrostatic interactions. In confirmation of this conclusion, replacement of the K at

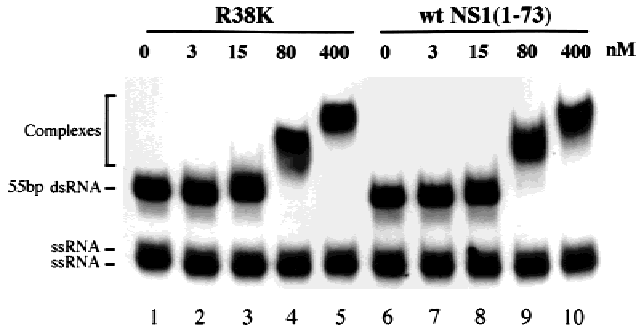


FIGURE 6. The interaction between RNA and the basic amino acid at position 38 of the NS1 RNA-binding domain is based on electrostatic interactions. The NS1A(1–73) polypeptide containing K instead of R at position 38 is designated as R38K. Increasing amounts of the R38K (lanes 2–5) or the wild-type (lanes 7–10) NS1(1–73) polypeptides were incubated with the 55-bp dsRNA target. Gel shift assays were performed as described previously. Lanes 1 and 6: dsRNA alone.

position 41 with a R does not reduce the affinity of the resulting NS1A(1–73) molecule for binding to either the 55 or the 140-bp dsRNA (data not shown).

An intact helix 3 is required for the RNA-binding and dimerization activities of NS1A(1–73)

Although an amino-terminal fragment of the NS1A protein containing 62, rather than 73, amino acids fails to bind U6 snRNA (Qian et al., 1995), alanine replacements of the lysine residues at positions 67 and 70 in helix 3 do not affect RNA binding or dimerization (Fig. 3). To determine whether the amino acid sequence containing these lysine residues serves any role, an amino-terminal fragment containing only the amino-terminal 65 amino acids (NS1A(1–65)) was constructed. This fragment does not bind dsRNA (Fig. 7A), and does not form a heterodimer with wild-type NS1A protein (Figure 7B). Consequently, an intact helix 3 (amino acids 54–69) is required for the stability of the NS1A(1–73) dimer, and hence for its RNA-binding activity. In fact, the three-dimensional structure of NS1A (1–73) shows that there are strong hydrophobic interactions between residues in helix 3 and residues in other helices in the dimeric structure (Chien et al., 1997; Liu et al., 1997). These interactions appear to be crucial for stabilizing this dimeric structure.

Basic residues in helix 2 that are required for, or enhance, RNA binding are conserved between the NS1 proteins of influenza A and B viruses

The NS1 protein of influenza B virus (NS1B protein) is larger than, and has very low sequence homology with, the NS1 protein of influenza A virus (NS1A protein; Briedis & Lamb, 1982). Nonetheless, the NS1B protein has the same RNA-binding activities as the NS1A protein (Wang & Krug, 1996). In addition, as is the case with the NS1A protein, an amino-terminal fragment of the NS1B protein possesses all the RNA-binding and dimerization activities of the full-length protein (Wang & Krug, 1996). The functional amino-terminal fragment of the NS1B protein is 93 amino acids long, and is thus about 20 amino acids longer than the functional amino-terminal fragment of the NS1A protein. Amide circular dichroism spectra showed that NS1B(1–93) is also largely α-helical (unpubl. data), suggesting that the RNA-binding and dimerization domains of NS1A and NS1B proteins share similar three-dimensional structures.

Accordingly, we aligned the amino-acid sequences of the NS1A(1–73) and NS1B(1–93) polypeptides, using primarily the Bestfit program in GCG (Fig. 8). The sequence similarity between the two sequences is about 40%, and sequence identity is only 20%. Nonetheless, a reasonable alignment was possible. This alignment suggested that the NS1B(1–93) polypeptide is proba-

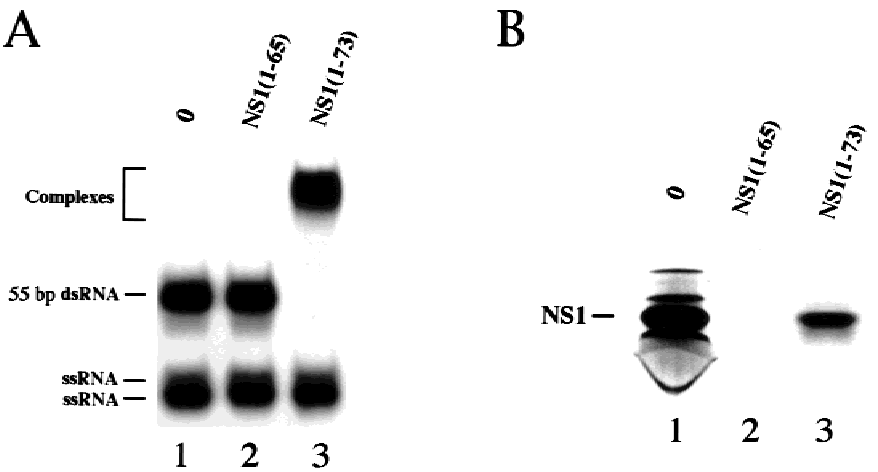


FIGURE 7. The entirety of helix 3 is required for the RNA-binding and dimerization activities of NS1A(1–73). Gel shift assays (**A**) and glutathione Sepharose selection assays (**B**) were performed in the absence (lane 1) or presence (400 nM) of the indicated amino-terminal NS1A polypeptides.

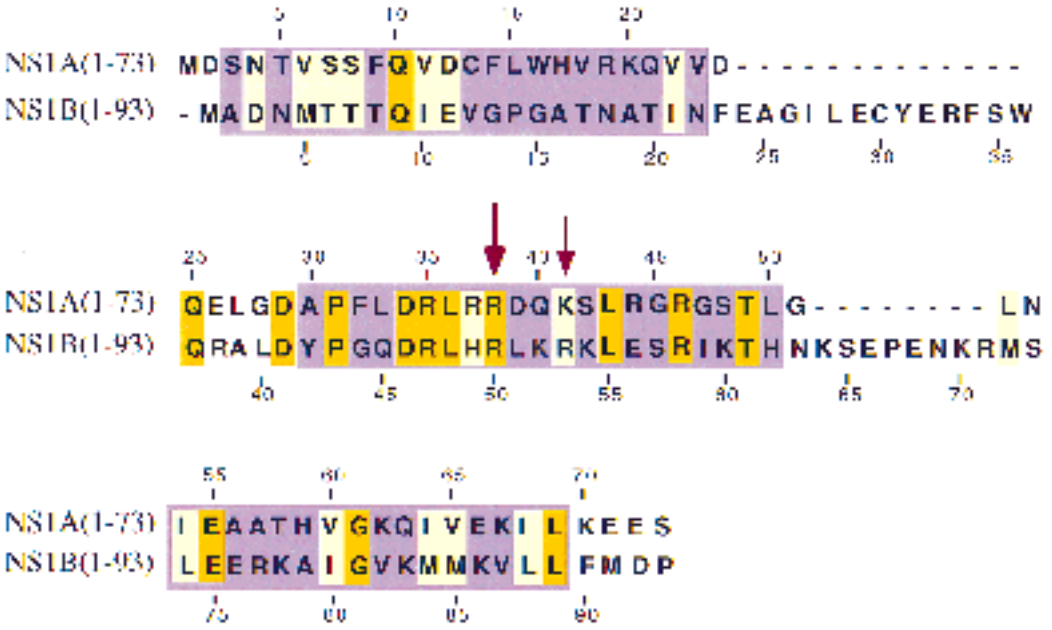


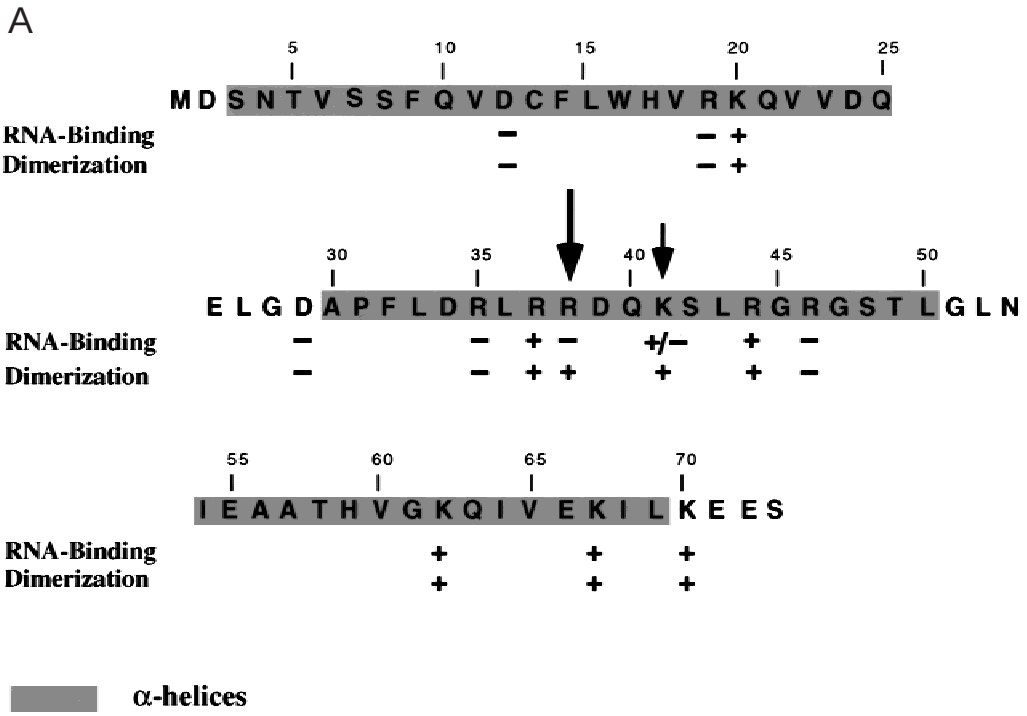
FIGURE 8. Sequence alignment of the RNA-binding/dimerization domains of NS1A (A/Udorn/72) and NS1B (B/Lee/40) proteins. The purple color denotes the three α-helices of NS1A(1–73). Identical residues are marked by orange, and conserved residues are marked by yellow. The dashes in the sequences indicate gaps. Red arrows indicate the two basic residues in helix 2 that are required for, or enhance, the RNA binding of NS1A and that are conserved in NS1B.

bly also comprised of three α-helices, with all the additional residues in NS1B(1–93) appearing in the loop regions between the three helices. The most conserved region is in helix 2, corresponding to amino acids 30–50 for NS1A and amino acids 43–62 for NS1B. Five (R35, R37, R38, K41, and R46) out of the six basic residues in helix 2 of NS1A(1–73) have counterparts in helix 2 of NS1B(1–93): R47, H49, R50, R53, and R58. In particular, the two basic residues in the NS1A(1–73) molecule (R38 and K41) that are required for, or strongly enhance, RNA binding are preserved in the NS1B(1–93) molecule (R50 and R53), suggesting that this is the essential region for RNA recognition in the NS1B protein as well as in the NS1A protein. However, the R44 basic residue in helix 2 of NS1A is not conserved as a basic residue in helix 2 of NS1B RNA. In contrast to helix 2, the basic residues in helix 3 are not conserved. Instead, helix 3 exhibits a conservation of some key buried hydrophobic residues, consistent with the hypothesis that the hydrophobic interactions in this helix contribute to the stability of the dimeric structure of the RNA-binding domain of both the NS1A and NS1B proteins.

DISCUSSION

The RNA-binding/dimerization domain of influenza A virus NS1 protein exhibits a novel dimeric six-helical fold (Chien et al., 1997; Liu et al., 1997). Based on the structure of NS1A(1–73), we carried out site-directed mutagenesis to introduce single alanine replacements to determine how this complex α-helical domain interacts with its RNA targets. The RNA-binding and dimerization activities of all the NS1A(1–73) alanine-scanning mutants are summarized along the amino-acid sequence in Figure 9A. The results demonstrate that the dimer structure is essential for RNA-binding activity, because all of the mutants that lack dimerization activity also lack RNA-binding activity, and all mutants that retain RNA-binding activity also retain dimerization activity.

In addition, we identified a few amino acids that affected only RNA binding and not dimerization. Of these amino acids, only one, the arginine at position 38 in helix 2, is absolutely required for RNA binding under all of the conditions studied, and thus its side chain is proposed to interact directly with the RNA target. This interaction is primarily electrostatic, because replacement of this arginine with lysine has no effect on RNA binding. A second basic amino acid, the lysine at position 41 in helix 2, makes an additional strong contribution to the affinity of RNA binding. Thus, in the absence of this lysine side chain, the affinity of binding to a high-affinity RNA substrate is reduced tenfold, and binding to a low-affinity RNA substrate is undetectable. In confirmation of the importance of these two basic residues, our alignment of the sequences of the NS1A and NS1B proteins indicates that basic amino acids are also present at the two comparable positions of the putative helix 2 of the NS1B protein. Our results also suggest that two other arginines at positions 37 and 44 in helix 2 of the NS1A(1–73) may contribute to a lesser extent to RNA binding. Thus, in the absence of either of



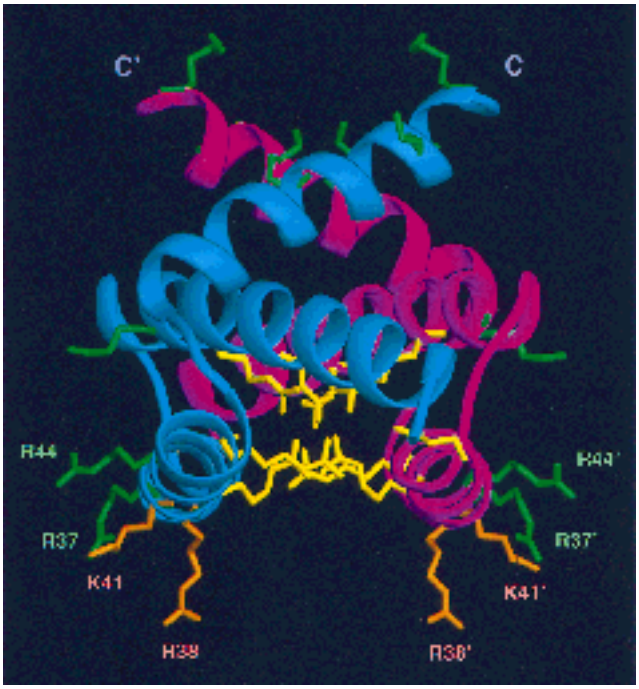


FIGURE 9. **A**: Summary of the phenotypes of the NS1A(1–73) polypeptides containing alanine replacements at the indicated positions. **B**: The RNA-binding epitope depicted on the three-dimensional structure of the dimeric NS1(1–73) molecule. The four basic amino acids (R38, R38′, K41, K41′)that are required for, or strongly enhance, RNA binding are shown in orange. All the other solvent-exposed basic amino acids are shown in green. The amino acids that have been shown to stabilize the dimer structure are shown in yellow.

these side chains, binding to a high-affinity RNA substrate is indistinguishable from that observed with a wild-type molecule, whereas binding to a low-affinity RNA substrate yields complexes with aberrant gel electrophoretic mobilities, specifically mobilities that differ only slightly from the free RNA substrate. The aberrant gel mobility of these complexes may indicate that they differ significantly in structure and/or electrostatics from complexes containing wild-type NS1A(1–73), a conclusion that can also be drawn from our observation that, unlike wild-type NS1A(1–73), the complexes formed between dsRNA and the R37A polypeptide decrease in mobility as the concentration of this polypeptide is increased. Perhaps R37 plays a more important role in RNA binding than R44, because a basic amino acid (His) is present only at the comparable position of R37 and not at the position of R44 in the proposed helix 2 of the NS1B protein.

Based on these results, we conclude that helix 2 and helix 2′, which are antiparallel and next to each other in

the dimer conformation, constitute the interaction face between the NS1A RNA-binding domain and its RNA targets (Fig. 9B). NMR chemical shift perturbation data also identifies this region of NS1A(1–73) as the RNA-binding epitope (Chien et al., in prep.). Consequently, unlike the HIV-1 Rev model peptide that contains a single α-helix (Battiste et al., 1996), the dimeric six-helical NS1A RNA-binding domain almost certainly does not bind in the major groove of its RNA target.

Rather, we predict that the NS1A protein uses at least one basic amino acid (R38, R38′) in each of the antiparallel helices 2 and 2′ (shown in orange in Fig. 9B) to contact the negatively charged phosphate backbone of dsRNA. Consequently, the dimer structure serves to position R38 on helix 2 so that it is a specific distance (~16.5 Å) from the corresponding amino acid R38′ on helix 2′. This distance is close to that (~17 Å) between the antiparallel phosphodiester backbones surrounding the minor groove. A second basic amino acid (K41, K41′) in each of the antiparallel helices 2 and 2′ (also shown in orange) also plays a role in binding to dsRNA; it may also contact the phosphate backbone of dsRNA, or it may provide a proximal positive charge to facilitate the interaction of R38 and R38′ with the negatively charged RNA backbone. In addition, binding may be facilitated by the proximal positive charges provided by two other basic amino acids (R37, R37′, R44, R44′) in helices 2 and 2′. An attractive working model of this complex proposed previously (Chien et al., 1997) postulates that antiparallel helices 2 and 2′ of NS1A(1–73) are aligned with, and span, the minor groove of A form dsRNA with contacts between one or more basic amino acids in helices 2 and 2′ and the polyphosphate backbone of the RNA. An analogous situation has been observed in the crystal structure of a portion of human topoisomerase I complexed with dsDNA (Stewart et al., 1998). In this complex only two out of nine basic amino acids on one face of a DNA-binding helix have direct contacts with phosphate groups that span the minor groove of the DNA structure.

The NS1A protein recognizes multiple RNA targets: not only dsRNA, but also U6 snRNA, U6atac snRNA, and poly(A) (Qiu & Krug, 1994; Lu et al., 1995; Qiu et al., 1995; Wang & Krug, 1998). Our results indicate that the RNA-binding domain of the NS1A protein uses the same epitope to interact with dsRNA and U6 snRNA. In addition, poly(A) and U6atac snRNA also interact with the same epitope (unpubl. experiments). This raises the question of whether these different RNA targets share at least some common structural features that enable them to be recognized by the same protein epitope. For example, it is quite possible that the NS1A RNA-binding domain recognizes a particular type of double-stranded RNA conformation when it binds to the stem–bulge structure in U6 snRNA.

Several other protein domains have been shown to bind dsRNA in a sequence-independent way. A prevalent dsRNA-binding motif that is found in a relatively large number of proteins is a sequence of approximately 65 amino acids long with an α-β-β-β-α structure (Bycroft et al., 1995; Kharrat et al., 1995). In addition, at least one type of C_2H_2 zinc finger motif binds dsRNA (Finerty & Bass, 1997). Like the six-helical NS1 RNA-binding domain, these other protein motifs bind to the A form helix of dsRNA, which has a very deep and narrow major groove and a wide and shallow minor groove (Weeks & Crothers, 1993). Because the binding of these protein domains to dsRNA is sequence independent, it is likely that they do not distort the major groove to gain access to the functional groups that distinguish different base pairs. Rather, they likely bind to the phosphate backbone and/or in the minor groove of A form dsRNA. It will be of considerable interest to determine whether the mode of interaction of these other dsRNA-binding protein domains with the A form helix is similar to that of the six-helical NS1 RNA-binding domain.

The present results indicate that the interpretation of our previous mutagenesis experiments (Qian et al., 1994), which were carried out without the benefit of structural information, may need to be modified. In these experiments, groups of two or three amino acids at various positions along the NS1A protein chain were replaced with alanines (Qian et al., 1994). Alanine replacement of three such amino acid groups near the amino-terminus of the protein resulted in a loss of the RNA-binding activity of the NS1A protein and a loss of the ability of the NS1A protein to inhibit both pre-mRNA splicing and the nuclear export of mRNA in vivo (Lu et al., 1994; Qian et al., 1994; Qiu & Krug, 1994). These results were interpreted as indicating that the RNA-binding activity of the NS1A protein is required for these two in vivo functions. However, as shown in the present study, many of the alanines introduced in the earlier experiments replace amino acids that are required for the dimerization of the NS1A protein. The loss of dimerization caused by such alanine replacements is a profound change in the structure of the NS1A protein that could lead to inactivation of the other functional domain of the NS1A protein, the effector domain. Because we have now identified a small number of basic amino acids that participate only in RNA binding, and not in dimerization, we are now in a position to determine which in vivo functions of the NS1A protein truly require RNA binding per se.

MATERIALS AND METHODS

Preparation of GST-NS1(1–73) fusion proteins

The A/Udorn/72 NS1(1–73) gene was cloned into the unique *Bam*HI site of the pGEX-3X plasmid. Oligonucleotide-directed mutagenesis was used to introduce point mutations into NS1A(1–73) DNA. The mutants were cloned into the pGEX-3X

vector, and their sequence was confirmed by dideoxynucleotide sequencing. The recombinant pGEX-3X constructs were transformed into *E. coli* strain JM101, and the expressed GST fusion proteins were purified by affinity chromatography on a glutathione Sepharose 4B column. Where indicated, the GST portion was cleaved from the fusion protein by treatment with protease factor Xa (1:100 ratio w/w with respect to the fusion protein) for 16 h at 4 °C in 100 mM NaCl, 50 mM Tris-HCl (pH 8.0) and 1 mM $CaCl_2$. The purity of the purified GST fusion proteins and the completeness of cleavage were established by gel electrophoresis followed by Coomassie blue staining.

RNA-binding assays

The GST portion of all the GST fusion proteins used in RNA-binding assays was cleaved using the protease factor Xa. The 55-bp dsRNA and U6 snRNA were prepared as described previously (Lu et al., 1995; Qiu et al., 1995). The 140-bp dsRNA was prepared by annealing the sense and anti-sense transcripts of a 140-nt-long globin sequence inserted in the pGEM1 vector. The indicated amount of an NS1A(1–73) polypeptide was mixed with the in vitro-transcribed, labeled RNA targets (10,000 cpm, 1 nM) in an RNA-binding solution in a total volume of 20 μL. The RNA-binding solution contained 50 mM Tris-HCl (pH 8.8), 50 mM glycine, 8% glycerol, 2.0 mM dithiolthreitol, 50 ng/μL *E. coli* tRNA, and 40 U of RNasin. The mixture was incubated on ice for 30 min. The RNA–polypeptide complexes were resolved from free RNA by nondenaturing electrophoresis on a 6% polyacrylamide gel for 3 h at 4 °C at 150 V using 45 mM Tris-borate, 1 mM EDTA as running buffer.

Glutathione Sepharose affinity assay for dimerization

A purified GST-NS1A(1–73) fusion protein (5 μg) was combined with 10 μL of in vitro-translated ^{35}S-labeled full-length NS1A protein, 20 μL of glutathione Sepharose 4B beads, and 400 μL of IPP_{150} buffer (10 mM Tris-HCl (pH 8.0), 150 mM NaCl, and 0.1% Nonidet P-40) in 1.5-mL microfuge tubes. The tubes were incubated at 4 °C for 1.5 h with constant rotation. The beads were washed three times with 1 mL of IPP_{150} buffer, followed by a wash in IPP_{375} buffer (same as IPP_{150} buffer but with 375 mM NaCl). The beads were then resuspended in gel-loading buffer containing 5% (v/v) 2-mercaptoethanol and heated at 95 °C for 2 min. The resulting eluate was analyzed by gel electrophoresis on a 14% SDS-polyacrylamide gel. The ^{35}S-labeled proteins were detected by fluorography of the SDS-polyacrylamide gel.

ACKNOWLEDGMENTS

This investigation was supported by grant AI11772 (to R.M.K.) and GM47014 (to G.T.M.) from the U.S. National Institutes of Health.

Received September 9, 1998; returned for revision October 13, 1998; revised manuscript received October 27, 1998

REFERENCES

Alonso-Caplen FV, Nemeroff ME, Yun Q, Krug RM. 1992. Nucleocytoplasmic transport: The influenza virus NS1 protein regulates the transport of spliced NS2 mRNA and its precursor NS1 mRNA. *Genes & Dev 6*:255–267.

Banner DW, Kokkinidis M, Tsernoglou D. 1987. Structure of the ColE1 rop protein at 1.7 Å resolution. *J Mol Biol 196*:657–675.

Battiste J, Mao H, Rao N, Tan R, Muhandiram D, Kay L, Frankel A, Williamson J. 1996. Alpha helix-RNA major groove recognition in an HIV-1 rev peptide-RRE RNA complex. *Science 273*:1547–1551.

Berglund H, Rak A, Serganov A, Garber M, Hard T. 1997. Solution structure of the ribosomal RNA binding protein S15 from *Thermus thermophilus*. *Nature Struct Biol 4*:20–23.

Briedis D, Lamb RA. 1982. Influenza B virus genome: Sequences and structural organization of RNA segment 8 and the mRNAs coding for the NS1 and NS2 proteins. *J Virol 42*:186–193.

Bycroft M, Grunert S, Murzin AG, Proctor M, St Johnston D. 1995. NMR solution structure of a dsRNA binding domain from *Drosophila staufen* protein reveals homology to the N-terminal domain of ribosomal protein S5. *EMBO J 14*:3563–3571.

Chien C-Y, Tejero R, Huang Y, Zimmerman DE, Krug RM, Montelione GT. 1997. A novel RNA binding motif in influenza A virus non-structural protein 1. *Nature Struct Biol 4*:891–895.

Davies C, Gerstner RB, Draper DE, Ramakrishnan V, White SW. 1998. The crystal structure of ribosomal protein S4 reveals a two-domain molecule with an extensive RNA-binding surface: One domain shows homology to the ETS DNA-binding motif. *EMBO J 17*:4545–4558.

Finerty PJ Jr, Bass BL. 1997. A *Xenopus* zinc finger protein that specifically binds dsRNA and RNA–DNA hybrids. *J Mol Biol 27*:195–208.

Fortes P, Beloso A, Ortin J. 1994. Influenza virus NS1 protein inhibits pre-mRNA splicing and blocks mRNA nucleocytoplasmic transport. *EMBO J 13*:704–712.

Hatada E, Futada R. 1992. Binding of influenza A virus NS1 protein to dsRNA in vitro. *J Gen Virol 73*:3325–3329.

Hinck AP, Markus MA, Huang S, Grzesiek S, Kustonovich I, Draper DE, Torchia DA. 1997. The RNA binding domain of ribosomal protein L11: Three-dimensional structure of the RNA-bound form of the protein and its interaction with 23 S rRNA. *J Mol Biol 274*:101–113.

Kharrat A, Macias MJ, Gibson TJ, Nilges M, Pastore A. 1995. Structure of the dsRNA binding domain of *E. coli* RNase III. *EMBO J 14*:3572–3584.

Liu J, Lynch PA, Chien C-Y, Montelione GT, Krug RM, Berman HM. 1997. Crystal structure of the unique RNA-binding domain of the influenza virus NS1 protein. *Nature Struct Biol 4*:896–899.

Lu Y, Qian X, Krug RM. 1994. The influenza virus NS1 protein: A novel inhibitor of pre-mRNA splicing. *Genes & Dev 8*:1817–1828.

Lu Y, Wambach M, Katze MG, Krug RM. 1995. Binding of the influenza virus NS1 protein to double-stranded RNA inhibits the activation of the protein kinase that phosphorylates the eIF-2 translation initiation factor. *Virology 214*:222–228.

Markus MA, Gerstner RB, Draper DE, Torchia DA. 1998. The solution structure of ribosomal protein S4 Δ41 reveals two subdomains and a positively charged surface that may interact with RNA. *EMBO J 17*:4559–4571.

Nagai K. 1996. RNA–protein complexes. *Curr Opin Struct Biol 6*:53–61.

Nemeroff ME, Barabino SM, Li Y, Keller W, Krug RM. 1998. Influenza virus NS1 protein interacts with the cellular 30K subunit of CPSF and inhibits 3′ end formation of cellular pre-mRNAs. *Mol Cell 1*:991–1000.

Nemeroff ME, Qian X, Krug RM. 1995. The influenza virus NS1 protein forms multimers in vitro and in vivo. *Virology 212*:422–428.

Qian X, Alonso-Caplen F, Krug RM. 1994. Two functional domains of the influenza virus NS1 protein are required for regulation of nuclear export of mRNA. *J Virol 68*:2433–2441.

Qian X, Chien C, Lu Y, Montelione GT, Krug RM. 1995. An amino-terminal polypeptide fragment of the influenza virus NS1 protein possesses specific RNA binding activity and largely helical backbone structure. *RNA 1*:948–956.

Qiu Y, Krug RM. 1994. The influenza virus NS1 protein is a poly(A)-binding protein that inhibits nuclear export of mRNAs containing poly(A). *J Virol 68*:2425–2432.

Qiu Y, Nemeroff ME, Krug RM. 1995. The influenza virus NS1 protein binds to a specific region in human U6 snRNA and inhibits U6–U2 and U6–U4 snRNA interactions during splicing. *RNA 1*:304–316.

Stebbins CE, Borukhov S, Orlova M, Polyakov A, Goldfarb A, Darst SA. 1995. Crystal structure of the GreA transcript cleavage factor from *Escherichia coli. Nature 373*:636–640.

Stewart L, Redinbo MR, Qiu X, Hol WG, Champoux JJ. 1998. A model for the mechanism of human topoisomerase I. *Science 279*:1534–1541.

Wang W, Krug RM. 1996. The RNA-binding and effector domains are conserved to different extents among influenza A and B viruses. *Virology 223*:41–50.

Wang W, Krug RM. 1998. U6atac snRNA, the highly divergent counterpart of U6 snRNA, is the specific target that mediates inhibition of AT-AC splicing by the influenza virus NS1 protein. *RNA 4*:55–64.

Weeks KM, Crothers DM. 1993. Major groove accessibility of RNA. *Science 261*:1574–1577.

Xing Y, GuhaThakurta D, Draper DE. 1997. The RNA binding domain of ribosomal protein L11 is structurally similar to homeodomains. *Nature Struc Biol 4*:24–27.